MALAYA JOURNAL OF MATEMATIK
Malaya J. Mat. **11(S)**(2023), 82–91.
http://doi.org/10.26637/mjm11S/006

Combinatorics on words obtaining by k to k substitution and k to k exchange of a letter on modulo-recurrent words

Moussa Barro[1], **K. Ernest Bognini**[2*] **and Théodore Tapsoba**[1]

[1]*Département de Mathématiques et Informatique, Université Nazi BONI,01 BP 1091 Bobo-Dioulasso 01, Burkina Faso.*
[2]*Centre Universitaire de Kaya (CU-K), Université Joseph KI-ZERBO, 03 BP 7021 Ouagadougou, Burkina Faso.*

Received 16 June 2023; Accepted 22 August 2023

This paper is dedicated to the occasion of Professor Gaston M. N'Guérékata's 70th birthday

Abstract. We introduce two new concepts which are the k to k substitution and k to k exchange of a letter on infinite words. After studying the return words and the special factors of words obtaining by these applications on Sturmian words and modulo-recurrent words. Next, we establish the complexity functions of these words. Finally, we determine the palindromic complexity of these words.

AMS Subject Classifications: 68R15, 11B85, 03D15.

Keywords: Sturmian words, modulo-recursive, k to k substitution, k to k exchange, complexity function, palindrome.

Contents

1. Introduction

The complexity function, which counts the number of distinct factors of given length in some infinite word, is often used in characterization of some families of words [1]. For instance, Sturmian words are the infinite words non eventually periodic with minimal complexity [9, 10]. Over the past thirty years, Sturmian words are intensively studied. Thus, these investigations has led to numerous characterizations and various properties [5, 6, 8, 13, 14] on these words. In [2, 11, 12], the palindromic factors are used abundantly to study Sturmian words.

*Corresponding author. Email addresses: **mous.barro@yahoo.com** (Moussa Barro), **ernestk.bognini@yahoo.fr** (K. Ernest Bognini), **tmytapsoba@yahoo.fr** (Théodore Tapsoba)

The notion of k to k insertion of a letter on infinite words was introduced in [13], and widely studied in [4, 12]. Later in [3], authors has studied the notion of k to k erasure of letter on infinite words .

Now, we introduce the concepts of k to k substitution and k to k exchange of letter on infinite words. The first notion consists to substitute a letter steadily with step of length k in some infinite words namely Strumian words and modulo-recurrent words. Thus, the new word obtaining by this application is called *word by k to k substitution* of letter. The second consists to exchange a letter steadily with step of length k in the same infinite words. Therefore, the new word obtaining by this application is called *word by k to k exchange* of letter. This paper is focused in the studying of combinatorial properties of these two new types of words obtaining on strumian words and modulo-recurrent words.

The paper is organized as follow. In the Section 2, we give useful definitions and notations in combinatorics on words, and we recall some properties of Sturmian words and modulo-recurrent words. In Section 3, we study the special factors and we determine the complexity of words obtained by k to k substitution and by k to k exchange of letter on uniformly modulo-recurrent words. The study of some palindromic properties and the palindromic complexity of these words are established in Secttion 4.

2. Background

2.1. Combinatorial properties

An alphabet $\mathcal{A}$, is a non empty finite set whose the elements are called letters. A word is a finite or infinite sequence of elements over $\mathcal{A}$. The set of finite words over $\mathcal{A}$ is denoted $\mathcal{A}^*$ and ε, the empty word. For any $u \in \mathcal{A}^*$, the number of letters of u is called length of u and it is denoted $|u|$. Moreover, for any letter x of $\mathcal{A}$, $|u|_x$ is the number of occurrences of x in u. A word u of length n written with a unique letter x is simply denoted $u = x^n$.

Let $u = x_1x_2 \cdots x_n$ be a word such that $x_i \in \mathcal{A}$, for all $i \in \{1, 2, \ldots, n\}$. The image of u by the reversal map is the word denoted $\overline{u}$ and defined by $\overline{u} = x_n \cdots x_2x_1$. The word $\overline{u}$ is simply called reversal image of u. A finite word u is called palindrome if $\overline{u} = u$. If u and v are two finite words over $\mathcal{A}$, we have $\overline{uv}$=$\overline{v}\,\overline{u}$.

The set of infinite words over $\mathcal{A}$ is denoted $\mathcal{A}^\omega$ and we write $\mathcal{A}^\infty = \mathcal{A}^* \cup \mathcal{A}^\omega$, the set of finite and infinite words. An infinite word **u** is said to be aperiodic if there exist two words $v \in \mathcal{A}^*$ and $w \in \mathcal{A}^+$ such that $u = vw^\omega$. If $v = \varepsilon$, then u is periodic.

Let $\mathbf{u} \in \mathcal{A}^\infty$ and $w \in \mathcal{A}^*$. The word w is a factor of u if there exist $u_1 \in \mathcal{A}^*$ and $\mathbf{u}_2 \in \mathcal{A}^\infty$ such that $\mathbf{u} = u_1w\mathbf{u}_2$. The factor w is said to be a prefix (respectively, a suffix) if u_1 (respectively, $\mathbf{u}_2$) is the empty word.

A word **u** is said to be recurrent if each of its factors appears infinitly in **u**. A word **u** is said to be uniformly recurrent if for all integers n, there exists an integer N such that any factor of length N in **u** contains all the factors of length n.

A non-empty factor w of **u**, is said to be right (respectively, left) extendable by a letter x in **u** if wx (respectively, xw) appears in **u**. The number of right (respectively, left) extensions of w, is denoted ∂^+w (respectively, ∂^-w). The factor w is said to be right (respectively, left) special in **u** if $\partial^+w > 1$ (respectively, $\partial^-w > 1$). If $\partial^+w = 2$ (respectively, $\partial^+w = 3$), we say that w have two-right (respectively, three-right) extensions in **u**. The factor w is said to be bispecial in **u** if w is both right and left special.

Let **u** be an infinite word over $\mathcal{A}$. The set of factors of length n in **u**, is written $\mathrm{L}_n(\mathbf{u})$ and the set of all factors in **u** is denoted by $\mathrm{L}(\mathbf{u})$. Let $\mathbf{u} = x_0x_1x_2\cdots$, where $x_i \in \mathcal{A}$, $i \geq 0$ be an infinite word and w his factor. Then, w appears in **u** at the position l if $w = x_lx_{l+1}\cdots x_{l+n-1}$.

The complexity function of a given infinite word **u** is the map of $\mathbb{N}$ to $\mathbb{N}^*$ defined by $p_{\mathbf{u}}(n) = \#\mathrm{L}_n(\mathbf{u})$, where $\#\mathrm{L}_n(\mathbf{u})$ designates the cardinal of $\mathrm{L}_n(\mathbf{u})$.

This function is related to the special factors by the relation (see [7]):

$$p_{\mathbf{u}}(n+1) - p_{\mathbf{u}}(n) = \sum_{w \in \mathrm{L}_n(\mathbf{u})} (\partial^+(w) - 1).$$

We call first rigth (respectively, left) difference of the complexity function $p_{\mathbf{u}}$, the functions defined for any integer n by:

$$r_{\mathbf{u}}(n) = \sum_{w \in \mathrm{L}_n(\mathbf{u})} (\partial^+(w) - 1) \text{ and } l_{\mathbf{u}}(n) = \sum_{w \in \mathrm{L}_n(\mathbf{u})} (\partial^-(w) - 1).$$

The set of palindromic factors of length n in $\mathbf{u}$ is denoted $\mathrm{Pal}_n(\mathbf{u})$, and the set of all palindromic factors in $\mathbf{u}$, is denoted $\mathrm{Pal}(\mathbf{u})$. The palindromic complexity function of $\mathbf{u}$ is the map of $\mathbb{N}$ to $\mathbb{N}$, defined by:

$$p_{\mathbf{u}}^{al}(n) = \#\mathrm{Pal}_n(\mathbf{u}).$$

A word $\mathbf{u} \in \mathcal{A}^\infty$ is said to be rich if for any factor w of $\mathbf{u}$, the number of palindromic factor in w is $|w| + 1$.

Let $\mathbf{u} = x_0x_1x_2x_3\cdots$ be an infinite word. The window complexity function of $\mathbf{u}$ is the map, $p_{\mathbf{u}}^f$ of $\mathbb{N}$ into $\mathbb{N}^*$, defined by:

$$p_{\mathbf{u}}^f(n) = \#\left\{x_{kn}x_{kn+1}\cdots x_{n(k+1)-1} : k \geq 0\right\}.$$

The shift is the application S of $\mathcal{A}^\omega$ to $\mathcal{A}^\omega$ which erases the first letter of some given word. For instance, $S(\mathbf{u}) = x_1x_2\cdots$.

A morphism f is a map of $\mathcal{A}^*$ into itself such that $f(uv) = f(u)f(v)$, for any $u, v \in \mathcal{A}^*$.

2.2. Sturmian words and modulo-recurrent words

In this subsection, we recall some useful properties on Sturmian words and modulo-recurrent words.

Definition 2.1. *An infinite word* $\mathbf{u}$ *over* $\mathcal{A}_2 = \{a, b\}$ *is said to be Sturmian if for any integer* n*,* $p_{\mathbf{u}}(n) = n + 1$.

The well-known Sturmian word in the literature is the famous Fibonacci word. It is generated by the morphism φ defined over $\mathcal{A}_2 = \{a, b\}$ by:

$$\varphi(a) = ab \text{ and } \varphi(b) = a.$$

Definition 2.2. *[8] An infinite word* $\mathbf{u} = x_0x_1x_2\cdots$ *is said to be modulo-recurrent if, any factor* w *of* $\mathbf{u}$ *appears in* $\mathbf{u}$ *at all positions modulo* i*, for all* $i \geq 1$.

Proposition 2.1. *[8] Let* $\mathbf{u} \in \mathcal{A}^\infty$ *such that* $p_{\mathbf{u}}(n) = (\#\mathcal{A})^n$*, for all* $n \in \mathbb{N}$*. Then,* $\mathbf{u}$ *is a modulo-recurrent word.*

Definition 2.3. *Let* $\mathbf{u}$ *be a modulo-recurrent word. Then,* $\mathbf{u}$ *is called non-trivial if there exists an integer* n_0 *such that for all* $n \geq n_0$*:*

$$p_{\mathbf{u}}(n) < (\#\mathcal{A})^n.$$

Definition 2.4. *Let* $\mathbf{u}$ *be an infinite word. Then,* $\mathbf{u}$ *is said to be uniformly modulo-recurrent if it is uniformly recurrent and modulo-recurrent.*

Definition 2.5. *A factor* w *of some infinite word* $\mathbf{u}$ *is said to be a window factor when it appears in* $\mathbf{u}$ *at a mutiple position its length.*

Let us denote $\mathrm{L}_n^f(\boldsymbol{u})$, the set of n-window factors of u for a given factor of length n. Thus, his cardinal is $p_{\boldsymbol{u}}^f(n)$.

Theorem 2.6. *[13] Every Sturmian word is modulo-recurrent.*

It is clear that the Sturmian words are non-trivial and uniformly modulo-recurrent. The champernowne word is modulo-recurrent but does not uniformly recurrent.

The following theorem presents some characterization on Sturmian words.

Theorem 2.7. *[11] Let* **u** *be a Sturmian word. Then, for all $n \in \mathbb{N}$, we have:*

$$p^{al}_{\mathbf{u}}(n) = \begin{cases} 1 & if \quad n \quad is \quad even \\ 2 & otherwise. \end{cases}$$

The modulo-recurrent words can be characterized by their window complexity as follow:

Theorem 2.8. *[8] A recurrent word* **u** *is modulo-recurrent if and only if $p^f_{\mathbf{u}}(n) = p_{\mathbf{u}}(n)$, for all $n \geq 1$.*

3. k to k substitution and k to k exchange of letter in infinite words

We introduce two new concepts called k to k substitution and k to k exchange of a letter in infinite words with $k \geq 1$. We study the combinatorial properties of these words.

Definition 3.1. *Let* **u** *be an infinite word over $\mathcal{A}$ in the form:*

$$\mathbf{u} = x_0 m_0 x_1 m_1 x_2 m_2 x_3 m_3 \cdots x_i m_i \cdots$$

with $m_i \in L_k(\mathbf{u})$ et $x_i \in \mathcal{A}$, $i \in \mathbb{N}$.

Let us substitute the letters x_i with a letter $c \notin \mathcal{A}$ in **u***. Then, we obtain an infinite word:*

$$\mathbf{v} = c m_0 c m_1 c m_2 c m_3 c \cdots c m_i c \cdots .$$

This new word is called word by k to k substitution of letter in **u** *and is denoted $\mathbf{v} = \mathcal{S}^c_k(\mathbf{u})$.*

Now, Let us denote $\tilde{.}$, the circular exchange map of letter defined over $\mathcal{A}$ and exchange the letters x_i in **u***. Thus, we obtain the word:*

$$\mathbf{w} = \tilde{x_0} m_0 \tilde{x_1} m_1 \tilde{x_2} m_2 \tilde{x_3} m_3 \cdots \tilde{x_i} m_i \cdots .$$

It is called word by k to k exchange of a letter in **u** *and denoted by $\mathbf{w} = \mathcal{E}_k(\mathbf{u})$.*

Example 3.1. Let us consider the Fibonacci word

$$\mathbf{f} = abaababaabaababaababaabababaabaababaa \cdots .$$

Then, we have:

$$\mathcal{S}^c_3(\mathbf{f}) = cbaacabacbaacabacbabcabacaabcabacaa \cdots ,$$

called word by 3 to 3 substitution of the letter c in **f**.

$$\mathcal{E}_2(\mathbf{f}) = bbaaaabbabbaaaabbabbabbabbaaaabbabbaaaaa \cdots ,$$

called word by 2 to 2 exchange of letter in **f**.

3.1. Return words and special factors in v

We study the return words of the extrenal letter c and special factors in **v**.

Proposition 3.1. *Let* **u** *be a recurrent word such that $\mathbf{v} = \mathcal{S}^c_3(\mathbf{u})$. Then, we have:*

$$Ret_{\mathbf{v}}(c) \subset \{cm_i; m_i \in \mathrm{L}_k(\mathbf{u})\}.$$

Proof. We have $Ret_{\mathbf{v}}(c) = \{cS(w_i);\ w_i \in \mathrm{L}^f_{k+1}(\mathbf{u})\}$. Moreover, we have $\{S(w_i); w_i \in \mathrm{L}^f_{k+1}(\mathbf{u})\} \subset \mathrm{L}_k(\mathbf{u})$. As a result, $\{cS(w_i); w_i \in \mathrm{L}^f_{k+1}(\mathbf{u})\} \subset \{cm_i; m_i \in \mathrm{L}_k(\mathbf{u})\}$. ■

Corollary 3.1. *Let* **u** *be a modulo-recurrent word such that* $\mathbf{v} = \mathcal{S}_3^c(\mathbf{u})$. *Then:*

$$\#Ret_{\mathbf{v}}(c) = p_{\mathbf{u}}(k).$$

Proof. According to Proposition 3.1, we have $Ret_{\mathbf{v}}(c) \subset \{cm_i; m_i \in \mathrm{L}_k(\mathbf{u})\}$. Let us consider $v_1 \in \{cm_i; m_i \in \mathrm{L}_k(\mathbf{u})\}$, then there exists $m_{i_0} \in \mathrm{L}_k(\mathbf{u})$ such that $v_1 = cm_{i_0}$. Since **u** is modulo-recurrent, then $cm_{i_0} \in \mathrm{L}(\mathbf{v})$, i.e. $v_1 \in \mathrm{L}(\mathbf{v})$. Hence, $\{cm_i;\ m_i \in \mathrm{L}_k(\mathbf{u})\} \subset Ret_{\mathbf{v}}(c)$. So, $Ret_{\mathbf{v}}(c) = \{cm_i;\ m_i \in \mathrm{L}_k(\mathbf{u})\}$, i.e. $\#Ret_{\mathbf{v}}(c) = p_{\mathbf{u}}(k)$. ■

Lemma 3.1. *Let* **u** *be a Sturmian word and* w *a factor of* $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$. *Then,* w *admits three-right (respectively, three-left) extensions in* **v** *if and only if* $|w| < k$ *and* w *is right (respectively, left) special in* **u**.

Proof. Let us consider $w \in \mathrm{L}(\mathbf{v})$.

CS: Let us assume that $|w| < k$ such that w is a rigth special factor of **u**. Then, wa and wb are factors of $textbfu$. In addition, $|wa| = |wb| \leq k$. Thus $wa, wb \in \mathrm{L}(\mathbf{v})$ since **u** is modulo-recurrent. Therefore, **u** being modulo-recurrent then by Corollary 3.1, we have $wc \in \mathrm{L}(\mathbf{v})$. So, w have three-right extensions in **v**.

CN: Let us suppose that w have three-right extensions in **v**. We discuss over the length of w.

Case 1: $|w| < k$. Since $wc \in \mathrm{L}(\mathbf{v})$, then $w \in \mathrm{L}(\mathbf{u})$ and $wa, wb \in \mathrm{L}(\mathbf{v})$. Thus, we have $wa, wb \in \mathrm{L}(\mathbf{u})$.

Case 2: $|w| = k$. Then, w have three-right extensions in **v**, i.e. $w \in \mathrm{L}(\mathbf{u})$. In addition, any factor of length $(k+1)$ in **v** contains exactly one occurrence of the letter c. That implies $wa, wb \notin \mathrm{L}(\mathbf{v})$. This contradicts the fact that w have three-right extensions in **v**.

Case 3: $|w| > k$. Then, w is in the form $w = w_0 c w_1$ with $|w_1| = k$. Since w is right special in **v** then $w_1 \in \mathrm{L}(\mathbf{u})$. By a similar reasoning to the case 2, we have a contradiction.

Hence, in all cases, w have three-right extensions in **v** if $|w| < k$.

With the same arguments we can show that w have also three-left extensions in **v**. ■

Theorem 3.2. *Let* **u** *be a Sturmian word over* $\mathcal{A}_2$ *and* $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$. *Then, the number of three-right (respectively, three-left) extensions of length* n *in* **v** *is:*

$$\boldsymbol{T}rip_{\mathbf{v}}(n) = \begin{cases} 1 & if \quad n < k \\ 0 & otherwise. \end{cases}$$

Proof. Let us consider $w \in \mathrm{L}(\mathbf{v})$. Then, by Lemma 3.1, w have three-right (respectively, three-left) extensions in **v** if and only if w is right (respectively, left) special in **u** and $|w| < k$. Since **u** is Sturmian, then for each length n, **u** have only one right (respectively, left) special factor. Consequently, **v** produces three-right (respectively, three-left) extensions factor of length n if $n < k$ and neither otherwise. ■

Theorem 3.3. *Let* **u** *be a Sturmian word over* $\mathcal{A}_2$ *and* $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$. *Then, the number of two-right (respectively, two-left) extensions of length* n *in* **v** *is:*

$$\boldsymbol{B}ip_{\mathbf{v}}(n) = \begin{cases} 2n+1 & if \quad n < k \\ k+1 & otherwise. \end{cases}$$

Proof. Let us designate by $\mathcal{B}r_{\mathbf{v}}(n)$, the set of factors of length n having two-right extensions in **v** and denote w_n the right special factor of length n in **u**. It follows that:

- For $n \leq k$, we have:

$$\mathcal{B}r_{\mathbf{v}}(n) = \mathrm{L}_n(\mathbf{u}) \backslash \{w_n\} \bigcup \{t_0 c t_1 : t_0 x t_1 = w_n; |t_1| = 0, 1 \cdots, n-1\} \bigcup \{tc : t = w_{n-1}\}.$$

Consequently, $\mathbf{B}ip_{\mathbf{v}}(n) = 2n + 1$.

• For $n > k$, we have:

$$\mathcal{B}r_{\mathbf{v}}(n) = \{t_0ct_1ct_2c\cdots ct_q : t_0x_1t_1x_2t_2x_3\cdots x_qt_q = w_n\}\bigcup\{tc : t = w_{n-1}\}$$

with $|t_0| \leq k, |t_q| \leq k-1$ and $|t_i| = k, \forall i = 1, 2, \cdots, q-1$.

Hence, $\mathbf{B}ip_{\mathbf{v}}(n) = k+1$. ■

Remark 3.1. Let **u** be a Sturmian word over $\mathcal{A}_2$ and $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$. Then, any right (respectively, left) special factor of length n in **v** with ($n \geq k$), give two-right (respectively, two-left) extensions in **v**.

3.2. Complexity function of the infinite word v

In this subsection, we determine the complexity function for the word **v** obtaining by k to k substitutiton of letter in some infinite word **u**.

Let $u_1 \in \mathrm{L}(\mathbf{u})$ for a given infinite word **u**. Then, we denote by $\mathcal{F}_k^c(u_1)$, the set of words obtaining by substitution of the letter c in u_1, i.e.

$$\mathcal{F}_k^c(u_1) = \{t_0ct_1ct_2c\cdots ct_q : t_0x_1t_1x_2t_2x_3\cdots x_qt_q = u_1; |t_0|, |t_q| \leq k, |t_i| = k, \forall i = 1, 2, \ldots, q-1\}.$$

Proposition 3.2. *Let* **u** *be a modulo-recurrent word such that* $u_1 \in \mathrm{L}(\mathbf{u})$ *and* $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$. *Then:*

$$\mathcal{F}_k^c(u_1) \subset \mathrm{L}(\mathbf{v}).$$

Remark 3.2. Any factor of length n in **v** comes from a factor of length n in **u**.

$$\#\mathcal{F}_k^c(u_1) = \begin{cases} |u_1| & if \quad |u_1| \geq k \\ k+1 & otherwise. \end{cases}$$

Lemma 3.2. *Let us consider* $v_1, v_2 \in \mathrm{L}_n(\mathbf{v})$. *Then:*

$$||v_1|_c - |v_2|_c| \leq 1.$$

Proof. By Remark 3.2, the k to k substitution conserve the lengths of factors. ■

Proposition 3.3. *Let* **u** *be an infinite word such that* $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$. *Then, for any integer* $n > 1$, *we have:*

$$p_{\mathbf{v}}(n) \leq \begin{cases} (n+1)p_{\mathbf{u}}(n) - r_{\mathbf{u}}(n-1) - l_{\mathbf{u}}(n-1) & if \quad n \leq k \\ (k+1)p_{\mathbf{u}}(n) - r_{\mathbf{u}}(n-1) - l_{\mathbf{u}}(n-1) & if \quad n > k. \end{cases}$$

Proof. Firstly, the substitutions starting with the first letter of extensions to the left of the same special factor to the left give the same factor. The same applies to substitutions ending in the last letter extensions to the right of the same special factor to the right give the same factor. Thus, according to the values of k and n we have two cases.

Case 1 : $n \leq k$. Then, some factors of **u** of length n are also factors of **v**. Moreover, the other factors of **v** of length n are produced from factors of **u** of length n by substitution of one letter. Thus, we have:

$$\mathrm{L}_n(\mathbf{v}) \subseteq \mathrm{L}_n(\mathbf{u})\bigcup\{rcs; \quad rxs \in \mathrm{L}_n(\mathbf{u}), x \in \mathcal{A}\}.$$

Consequently, we have the following inequality: $p_{\mathbf{v}}(n) \leq p_{\mathbf{u}}(n) + np_{\mathbf{u}}(n) - r_{\mathbf{u}}(n-1) - l_{\mathbf{u}}(n-1)$, i.e. $p_{\mathbf{v}}(n) \leq (n+1)p_{\mathbf{u}}(n) - r_{\mathbf{u}}(n-1) - l_{\mathbf{u}}(n-1)$, for all $n \leq k$.

Case 2 : $n > k$. Then, any factor of length n in **v** contains at least one occurrence of the letter c. By Remark 3.2, any factor of length n in **u** prodruces at most $(k+1)$ factors of length n in **v**. Hence, $p_{\mathbf{v}}(n) \leq (k+1)p_{\mathbf{u}}(n) - r_{\mathbf{u}}(n-1) - l_{\mathbf{u}}(n-1)$, for all $n > k$. ■

Corollary 3.2. *Let* **u** *be a uniformly modulo-recurrent word and* $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$*. Then, the complexity function of* **v** *is given by:*

$$p_{\mathbf{v}}(n) = \begin{cases} \#(\mathcal{A}) + 1 & if \quad n = 1 \\ (n+1)p_{\mathbf{u}}(n) - r_{\mathbf{u}}(n-1) - l_{\mathbf{u}}(n-1) & if \quad 1 < n \leq k. \end{cases}$$

Proof. Since **u** being modulo-recurrent, then by Proposition 3.2, all substitutions of factors of **u** give the factors of **v**. By using the Proposition 3.3, we deduce
$p_{\mathbf{v}}(n) = (n+1)p_{\mathbf{u}}(n) - r_{\mathbf{u}}(n-1) - l_{\mathbf{u}}(n-1)$ if $1 < n \leq k$. ■

For the following, we need the definition below.

Definition 3.4. *Let* **u** *be a uniformly modulo-recurrent word and* u_1 *be a factor of* **u**. *Then* u_1 *is said to be sufficiently long if, it contains all the* $(k+1)$*-window factors of* **u**.

Now, we denote n_0, the minimal length of the sufficiently long factors of ***u***. The next Lemma allowed to determine the complexity function of **v**.

Lemma 3.3. *Let* **u** *be a uniformly modulo-recurrent word over* $\mathcal{A}$. *Let* $u_1,\ u_2$ *be two distinct sufficiently long factors of* **u** *such that* $S(u_1) \neq S(u_2)$ *with* $u_1x^{-1} \neq u_2y^{-1}$ *where* $x, y \in \mathcal{A}$. *Then, the words obtained by* k *to* k *substitution of letter in* u_1 *and* u_2 *are distinct in* $\mathcal{S}_k^c(\mathbf{u})$.

Proof. Let us consider $|S(u_1)| \geq n_0$ and $|S(u_2)| \geq n_0$. We have two cases.

Case 1 : $|u_1| \neq |u_2|$. Then u_1 and u_2 give distinct factors, since the k to k substitution preserve the lengths.

Case 2 : $|u_1| = |u_2|$. Then, let us assume that $S(u_1) \neq S(u_2)$ and $u_1x^{-1} \neq u_2y^{-1}$. Since ***u*** is a modulo-recurrent and $|u_1|, |u_2| \geq n_0$, then we have $\mathcal{F}_k^c(u_1) \cap \mathcal{F}_k^c(u_2) = \emptyset$. ■

Theorem 3.5. *Let* **u** *be a uniformly modulo-recurrent word and* $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$. *Then, the complexity function of* **v** *is given by:*

$$p_{\mathbf{v}}(n) = (k+1)p_{\mathbf{u}}(n) - r_{\mathbf{u}}(n-1) - l_{\mathbf{u}}(n-1), \textit{ for all } n \geq n_0.$$

Proof. Let us suppose that **u** is uniformly modulo-recurrent. Then, by Proposition 3.2, all substitutions of factors of **u** give the factors of **v**. According to Lemma 3.3, we obtain the existence of n_0. Thus, the Proposition 3.5 allows us to deduce:

$$p_{\mathbf{v}}(n) = (k+1)p_{\mathbf{u}}(n) - r_{\mathbf{u}}(n-1) - l_{\mathbf{u}}(n-1), \text{ for all } n \geq n_0.$$

■

Theorem 3.6. *Let* **u** *be a uniformly modulo-recurrent word and* $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$. *Then, the window complexity function of* **v** *satisfies:*

$$p_{\mathbf{v}}^f(n) = \begin{cases} p_{\mathbf{u}}(n-1) & if \quad n \equiv 0 \ \ mod\ (k+1) \\ p_{\mathbf{v}}(n) & otherwise. \end{cases}$$

Proof. Let us consider $n \in \mathbb{N}$ and **u** modulo-recurrent. Then, we have:

• If $n \equiv 0 \ \ mod\ (k+1)$, then let us put $n = q(k+1)$, $q \geq 1$. It follows that:
$\mathrm{L}_n^f(\mathbf{v}) = \{cm_1cm_2c\cdots cm_q :\ x_1m_1x_2\cdots x_qm_q \in \mathrm{L}_n^f(\mathbf{u});\ |m_i| = k;\ i = 1, \ldots, q\}$
$\mathrm{L}_n^f(\mathbf{v}) = \{cm_1cm_2c\cdots cm_q :\ m_1x_2\cdots x_qm_q \in \mathrm{L}_{n-1}(\mathbf{u});\ |m_i| = k;\ i = 1, \ldots, q\}$.
Thus, $\#\mathrm{L}_n^f(\mathbf{v}) = \#\mathrm{L}_{n-1}(\mathbf{u})$, i.e. $p_{\mathbf{v}}^f(n) = p_{\mathbf{u}}(n-1)$.

• Otherwise we have allways that $\#\mathrm{L}_n^f(\mathbf{v}) = \#\mathrm{L}_n^f(\mathbf{u}) = \#\mathrm{L}_n(\mathbf{u})$. ■

Corollary 3.3. *Let* **u** *be a Sturmian word and* $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$. *Then, we have:*

- *for* $n \le k$,

$$p_{\mathbf{v}}(n) = \begin{cases} 3 & if \quad n = 1 \\ n^2 + 2n - 1 & if \quad 1 < n \le k; \end{cases}$$

- *for* $n \ge n_0$,

$$p_{\mathbf{v}}(n) = kn + n + k - 1, \quad for\ all \quad n \ge n_0.$$

Proof. We have $\mathrm{L}_1(\mathbf{v}) = \{a, b, c\}$, so $p_{\mathbf{v}}(1) = 3$. According to Definition 2.1, we have $r_{\mathbf{u}}(n) = l_{\mathbf{u}}(n) = 1$, for all $n \in \mathbb{N}$. Hence, by Corollary 3.2 we have $p_{\mathbf{v}}(n) = n^2 + 2n - 1$ if $1 < n \le k$ and by using the Theorem 3.5, we deduce that $p_{\mathbf{v}}(n) = kn + n + k - 1$, for all $n \ge n_0$. ■

Corollary 3.4. *Let* **u** *be a non-trivial and uniformly modulo-recurrent word over* $\mathcal{A}$ *and* $\mathbf{v} = \mathcal{S}_k^x(\mathbf{u})$ *with* $x \in \mathcal{A}$. *Then, the complexity function of* **v** *satisfies:*

$$p_{\mathbf{v}}(n) = (k+1)p_{\mathbf{u}}(n) - r_{\mathbf{u}}(n-1) - l_{\mathbf{u}}(n-1), \text{ for all } n \ge n_0.$$

Proof. The k to k substitution of an internal and external letter in the modulo-recurrent words are the same for longer lengths. In addition, ***u*** being uniformly recurrent, thus by Lemma 3.3, we have the existence of n_0. Hence, by Theorem 3.5, we deduce the equalty. ■

3.3. Complexity function of the infinite word w

Here we are focus our study on the complexity function of the word **w**.

Proposition 3.4. *Let* **u** *be an infinite word and* $\mathbf{w} = \mathcal{E}_k(\mathbf{u})$. *Then, we have:*

$$p_{\mathbf{w}}(n) \le \begin{cases} (n+1)p_{\mathbf{u}}(n) & if \quad n \le k \\ (k+1)p_{\mathbf{u}}(n) & if \quad n > k. \end{cases}$$

Proof. The demonstration is similar to Proposition 3.3. ■

Theorem 3.7. *Let* **u** *be a non-trivial and uniformly modulo-recurrent word and* $\mathbf{w} = \mathcal{E}_k(\mathbf{u})$. *Then, the complexity function of* **w** *is given by:*

$$p_{\mathbf{w}}(n) = (k+1)p_{\mathbf{u}}(n), \text{ for all } \ n \ge n_0.$$

Proof. Since **u** is uniformly modulo-recurrent. Then, by Definition 3.4, any factor of length n ($n \ge n_0$) in **u** produces $(k+1)$ distinct factors of length n in **w**. Hence, $p_{\mathbf{w}}(n) = (k+1)p_{\mathbf{u}}(n)$. ■

Remark 3.3. Let $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$ and $\mathbf{w} = \mathcal{E}_k(\mathbf{u})$ be two infinite words for a given word **u**. Then, we have:

(i) The word **v** (respectively, **w**) is aperiodic if and only if **u** is aperiodic.

(ii) The word **v** (respectively, **w**) is recurrent if and only if **u** is recurrent.

4. Palindromic study of v and w

In this section, we prove that **v** and **w** have the same palindromic complexity function if **u** is non-trivial and uniformly modulo-recurrent and then we give this complexity.

Proposition 4.1. *Let* **u** *be a modulo-recurrent word such that* $\mathbf{v} = \mathcal{S}_k^x(\mathbf{u})$ *and* $\mathbf{w} = \mathcal{E}_k(\mathbf{u})$ *with* $x \in \mathcal{A}$. *Then, the language* $\mathrm{L}(\mathbf{v})$ *(respectively,* $\mathrm{L}(\mathbf{w})$*) is stable by reversal map if and only if* $\mathrm{L}(\mathbf{u})$ *is stable by reversal map.*

Remark 4.1. Any palindromic factor of **v** (respectively, **w**) comes from a palindromic factor of ***u***.

Proposition 4.2. *Let* **u** *be an infinite word. Then, the palindromic complexity function of* **v** *is given by:*

- *for $n \leq k$,*

$$p^{al}_{\mathbf{v}}(n) \leq \begin{cases} p^{al}_{\mathbf{u}}(n) & if \quad n \ even \\ 2p^{al}_{\mathbf{u}}(n) & otherwise. \end{cases}$$

- *for $n > k$, there exists two integers q and r such that $n = (k+1)q + r$ with $0 \leq r \leq k$ and $q \geq 1$;*

$$p^{al}_{\mathbf{v}}(n) \leq \begin{cases} 2p^{al}_{\mathbf{u}}(n) & if \quad k, r \ odd \\ 0 & if \quad k \ odd, r \ even \\ p^{al}_{\mathbf{u}}(n) & otherwise. \end{cases}$$

Proof. • For $n \leq k$, the factors of length n in **v** are factors of **u** and those of **v** in the form rcs with $rxs \in \mathrm{L}_n(\mathbf{u})$. We obtain:

$$\mathrm{Pal}_n(\mathbf{v}) \subseteq \mathrm{Pal}_n(\mathbf{u}) \bigcup \left\{ tc\bar{t}; \quad tx\bar{t} \in \mathrm{Pal}_n(\mathbf{u}), x \in \mathcal{A}_2 \right\}.$$

- For $n > k$, we have $n = (k+1)q + r$ with $0 \leq r \leq k$, $q \geq 1$. According to Lemma 3.2, for $v_1 \in L_n(\mathbf{v})$ we have $|v_1|_c \in \{q, q+1\}$ with $q = \dfrac{n-r}{k+1} = \left\lfloor \dfrac{n}{k+1} \right\rfloor$.

  $$\mathrm{Pal}_n(\mathbf{v}) \subseteq \left\{ m_0cm_1cm_2c\cdots cm_q : m_0x_1m_1x_2\cdots x_qm_q \in \mathrm{Pal}_{(k+1)q+r}(\mathbf{u}); |m_0| = |m_q| = \frac{k+r}{2} \right\}$$

  $$\bigcup \left\{ m_0cm_1c\cdots cm_qcm_{q+1} : m_0x_1m_1\cdots x_{q+1}m_{q+1} \in \mathrm{Pal}_{(k+1)q+r}(\mathbf{u}); |m_0| = |m_{q+1}| = \frac{r-1}{2} \right\}.$$

  - If k and r are odd, then $k+r$ and $r-1$ are even. Thus, we have $p^{al}_{\mathbf{v}}(n) \leq 2p^{al}_{\mathbf{u}}(n)$.
  - If k and r are even, then $k+r$ is even and $r-1$ is odd. So, we have $p^{al}_{\mathbf{v}}(n) \leq p^{al}_{\mathbf{u}}(n)$.
  - If k is even and r is odd, then $k+r$ is odd and $r-1$ is even. Thus, we deduce $p^{al}_{\mathbf{v}}(n) \leq p^{al}_{\mathbf{u}}(n)$.
  - If k is odd and r is even, then $k+r$ and $r-1$ are odd. Consequently, **v** does not admit palindromic factor of length n.

■

Corollary 4.1. *Let* **u** *be a uniformly modulo-recurrent word over $\mathcal{A}$ such that $awa, bwb \in \mathrm{Pal}(\mathbf{u})$ implies $a = b$, for all $a, b \in \mathcal{A}$. Then, the inequality of the Proposition 4.2 becomes an eqality, for all $n > n_0$.*

Remark 4.2. Let **u** be a non-trivial and uniformly modulo-recurrent word over $\mathcal{A}$ such that $awa, bwb \in \mathrm{Pal}(\mathbf{u})$ implies $a = b$, for all $a, b \in \mathcal{A}$. Then, $\mathbf{v} = \mathcal{S}^x_k(\mathbf{u})$ and $\mathbf{w} = \mathcal{E}_k(\mathbf{u})$ verifie:

$$p^{al}_{\mathbf{w}}(n) = p^{al}_{\mathbf{v}}(n), \text{ for all } n > n_0.$$

Remark 4.3. The k to k substitution and k to k exchange do not preserve:

(i) the palindromic richness with exception where $k = 1$;

(ii) the modulo-recursive.

Theorem 4.1. *Let* **u** *be a Sturmian word over $\mathcal{A}_2$ and $\mathbf{v} = \mathcal{S}^c_k(\mathbf{u})$. Then, we have:*

- *for $n \leq k$,*

$$p^{al}_{\mathbf{v}}(n) = \begin{cases} 3 & if \quad n = 1 \\ 1 & if \quad n \ even \\ 4 & otherwise \end{cases}.$$

- *for* $n > n_0$,

$$p_{\mathbf{v}}^{al}(n) = \begin{cases} 4 & if \quad k, n\ odd \\ 2 & if \quad k\ even, n\ odd \\ 1 & if \quad k, n\ even \\ 0 & if \quad k\ odd, n\ even. \end{cases}$$

Proof. • For $n \leq k$. Then for the initial rank $n = 1$, we have $\mathrm{Pal}_1(\mathbf{v}) = \{a, b, c\}$.
Since **u** is Sturmian then, it is modulo-recurrent. So, by Theorem 3.5, for $1 < n \leq k$, we have:

$$\mathrm{Pal}_n(\mathbf{v}) = \mathrm{Pal}_n(\mathbf{u}) \bigcup \left\{ tc\bar{t}; \quad tx\bar{t} \in \mathrm{Pal}_n(\mathbf{u}), x \in \mathcal{A}_2 \right\}.$$

By Theorem 2.7, **v** admits 1 (respectively, 4) palindromic factors if n is even (respectively, odd).

- $n > k$, we write $n = (k+1)q + r$ with $0 \leq r \leq k$. By using the parity of k and n in Proposition 4.2 and Corollary 4.1, we have the result by Theorem 2.7 .

■

Remark 4.4. If **u** is Sturmian then $\mathbf{v} = \mathcal{S}_k^c(\mathbf{u})$ is not erasing Sturmian.

References

[1] J. -P. Allouche, Sur la complexité des suites infinies, *Bull. Belg. Math. Soc. Simon Stevin,* **1**(1994), 133–143.

[2] J. -P. Allouche, M. Baake, J. Cassaigne and D. Damanik, Palindrome complexity, *Theoret. Comput. Sci.*, **292**(2003), 9–33.

[3] M. Barro, I. Kaboré, T. Tapsoba, About the words by k to k erasure of letter and the words of erased letters: Sturmian case, *IJAM.*, **32(2)**(2019), 189–204.

[4] M. Barro, I. Kaboré, T. Tapsoba, On the words by k to k insertion of a letter in sturmian words, *IJAM.*, **30(5)**(2017), 387–400.

[5] J. Berstel, Sturmian and episturmian words (a survey of some recent resultats), *Lect. Notes in Comput. Sci.*, **4728**(2007).

[6] V. Berthé, Fréquences des facteurs des suites sturmiennes, *Theoret. Comput. Sci.*, **165**(1996), 295–309.

[7] J. Cassaigne, Complexité et facteurs spéciaux, *Bull. Belg. Math. Soc. Simon Stevin*, **4**(1997), 67–88.

[8] J. Cassaigne, I. Kaboré, T. Tapsoba, On a new notion of complexity on infinite words, *Acta Univ. Sapentiae, Mathematica,* **2**(2010), 127–136.

[9] E. M. Coven, Sequences with minimal block growth, *Math. Syst. Theory,* **8**(1975), 376–382.

[10] E. M. Coven, G. A. Hedlund, Sequences with minimal block growth, *Math. Syst. Theory*,**7**(1971),138–153.

[11] X. Droubay, G. Pirillo, Palindromes and Sturmian words, *Theoret. Comput. Sci.*, **223(1-2)**(1999), 73–85.

[12] I. Kaboré, About k to k insertion words of sturmian words, *IJPAM*, **79(4)**(2012), 561–572.

[13] I. Kaboré, T. Tapsoba,Combinatoire des mots récurrents de complexités $n + 2$, *RAIRO-Theo. Inf. Appl.*, **41**(2007), 425–446.

[14] M. Lothaire, Algebraic combinatorics on words, Cambridge University Press, 2002.